The Astronomical Journal, 141:173 (12pp), 2011 May doi:10.1088/0004-6256/141/5/173

ABSOLUTE FLUX CALIBRATION OF THE IRAC INSTRUMENT ON THE *SPITZER SPACE TELESCOPE* USING *HUBBLE SPACE TELESCOPE* FLUX STANDARDS

R. C. Bohlin[1], K. D. Gordon[1], G. H. Rieke[2], D. Ardila[3], S. Carey[4], S. Deustua[1], C. Engelbracht[2], H. C. Ferguson[1], K. Flanagan[1], J. Kalirai[1], M. Meixner[1], A. Noriega-Crespo[4], K. Y. L. Su[2], and P.-E. Tremblay[5]
[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[2] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA
[3] Caltech, NASA Herschel Science Center, IPAC/MS 100-22, Pasadena, CA 91125, USA
[4] Caltech, Spitzer Science Center, MS 220-6, Pasadena, CA 91125, USA
[5] Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, Québec H3C 3J7, Canada
Received 2011 February 10; accepted 2011 March 7; published 2011 April 7

ABSTRACT

The absolute flux calibration of the *James Webb Space Telescope* (*JWST*) will be based on a set of stars observed by the *Hubble* and *Spitzer Space Telescopes*. In order to cross-calibrate the two facilities, several A, G, and white dwarf stars are observed with both *Spitzer* and *Hubble* and are the prototypes for a set of *JWST* calibration standards. The flux calibration constants for the four *Spitzer* IRAC bands 1–4 are derived from these stars and are 2.3%, 1.9%, 2.0%, and 0.5% lower than the official cold-mission IRAC calibration of Reach et al., i.e., in agreement within their estimated errors of ∼2%. The causes of these differences lie primarily in the IRAC data reduction and secondarily in the spectral energy distributions of our standard stars. The independent IRAC 8 μm band-4 fluxes of Rieke et al. are about 1.5% ± 2% higher than those of Reach et al. and are also in agreement with our 8 μm result.

Key words: stars: atmospheres – stars: fundamental parameters – techniques: spectroscopic

1. INTRODUCTION

Flux calibrations in physical units for astronomical instruments are required to make comparisons to physical models of observed objects (Kent et al. 2009). In particular, one of the main incentives for accurate absolute flux standards is the requirement for measuring the relative fluxes of redshifted SN Ia spectra in the rest frame in order to constrain the parameters of the dark energy. These constraints depend only on the precision of the ratio of fluxes from one wavelength to another and not on the absolute flux level. Quantitative descriptions of dark energy are significantly improved when the relative flux with wavelength is known to an accuracy of 1% or better (Aldering et al. 2004).

Absolute flux calibrations of spectrometers and photometers are normally derived from observations of standard stars with well-known spectral energy distributions (SEDs). For all *Hubble Space Telescope* (*HST*) instruments, all flux calibrations are traceable to three primary white dwarf (WD) standards, G191B2B, GD71, and GD153. The slopes of these WD SEDs are determined by non-local thermodynamic equilibrium (NLTE) model calculations using the Hubeny Tlusty Version-203 code for pure hydrogen atmospheres (Bohlin 2000; Bohlin et al. 2001; Hubeny & Lanz 1995; Tremblay & Bergeron 2009). The effective temperature and gravity are determined by fitting the models to ground-based observations of the Balmer line profiles (Finley et al. 1997).

The absolute flux of the models for these three primary standards is set by Space Telescope Imaging Spectrograph (STIS) relative spectrophotometry (Bohlin & Gilliland 2004; Bohlin 2007a) and the Megessier (1995) absolute flux for Vega of 3.46×10^{-9} erg cm^{-2} s^{-1} Å^{-1} at 5556 Å (3560 Jy or 3562 Jy for vacuum wavelengths). As discussed in the review by Hayes (1985), by Megessier (1995), and in Section 4.1.1 below, there is a small uncertainty in the 5556 Å flux, but this uncertainty just affects the overall level and *not* the slope (i.e., "color") of the WD models used for *HST* flux calibrations. Despite suggestions of variations, Hayes (1985) discusses the evidence for variability of Vega and concludes that the star is likely not variable. However, Engelke et al. (2010) present evidence for a 0.08 mag variation of Vega at visible wavelengths.

To compare with previously published calibrations of the *Spitzer Space Telescope* in the four IRAC bands (Fazio et al. 2004; Reach et al. 2005, hereafter Re05), a set of new observations of WD, A stars, and solar-analog G stars were made near the end of the cold mission. *Spitzer* data in the IRS blue peakup channel (Houck et al. 2004) or the MIPS 24 μm band (Rieke et al. 2004; Engelbracht et al. 2007) were included to ensure that debris disks or red companions do not contaminate the results. These new observations are supplemented by more data sets for the same stars from the *Spitzer* archive. Table 1 lists the 14 stars with *HST*-based SEDs that are used for the comparison with the absolute flux calibrations of Re05 for IRAC, the IRS Instrument Handbook,[6] and Engelbracht et al. (2007) for the MIPS 24 μm channel. The *HST* flux distributions are all in the CALSPEC[7] database. The Re05 IRAC calibration is based on four A star SEDs from an extension of the original Cohen CWW network (Cohen et al. 1992b, 1999, 2003; Cohen 2007), as validated by Price et al. (2004). Two of these four primary A star calibrators, HD 165459 and 1812095, are included in Table 1, while our other five A stars are listed as IRAC candidate primary calibrators by Re05. For a comparison of the *HST* SEDs with the Cohen flux distributions, see Bohlin & Cohen (2008), whose minor revisions include average fluxes that are ∼0.5% lower in the IRAC wavelength range for the set of seven A stars in Table 1. The K star calibrators of Re05 are not utilized, because of the extra complexity of modeling the molecular absorption and because Re05 used only A stars to define their final IRAC calibration constants.

In this paper, Section 2 covers the fundamental equations and concept of photometric flux calibrations, while Section 3

[6] http://ssc.spitzer.caltech.edu/irs/irsinstrumenthandbook/IRS_Instrument_Handbook.pdf
[7] http://www.stsci.edu/hst/observatory/cdbs/calspec.html/

Table 1
HST and *Spitzer* Comparison Stars

Star	R.A. J2000	Decl. J2000	V (mag)	Sp.T	$T_{\rm eff}$ (K)
GD71	5 52 27.51	+15 53 16.6	13.032	DA1	32747
GD153	12 57 02.37	+22 01 56.0	13.346	DA1	38686
G191B2B	5 05 30.62	+52 49 54.0	11.781	DA0	61193
LDS749B	21 32 16.24	+00 15 14.4	14.674	DBQ4	13575
HD 165459	18 02 30.74	58 37 38.1	6.864	A4V	8600
1732526	17 32 52.64	+71 04 43.1	12.530	A4V	8500
1740346	17 40 34.7	65 27 15.0	12.478	A6V	8050
1743045	17 43 04.5	66 55 01.7	13.525	A8II	7650
1802271	18 02 27.17	+60 43 35.6	11.985	A2V	9100
1805292	18 05 29.3	64 27 52.1	12.278	A4V	8400
1812095	18 12 09.57	63 29 42.3	11.736	A5V	8250
HD 209458	22 03 10.78	+18 53 03.7	7.65	G0V	6080
P041C	14 51 58.19	+71 43 17.3	12.01	G0V	5960
P330E	16 31 33.85	+30 08 47.1	13.01	G0V	5820

compares the published calibration constants for four IRAC imaging modes to those that are derived from the *HST*-based SEDs. Section 4 compares our results with the IRAC calibrations of Re05 and with the 8 μm fluxes of Rieke et al. (2008, hereafter Ri08). Section 5 includes suggestions for future efforts to improve the flux calibration accuracy. Finally in the Appendix, the *HST* method of photometric flux calibration is illustrated for six *Spitzer* modes. The Appendix is not absolutely essential to the main thrust of this paper but does expand on several points as forward referenced in the main body. In addition, the Appendix attempts to provide a cohesive mathematical foundation for the student or practitioner of the art of flux calibration.

2. GENERIC CALIBRATION CONSTANTS

2.1. Equations

This section compares the *Spitzer* flux calibration derived from the *HST* flux standard stars using the Re05 methodology and nomenclature. For comparison, the Appendix presents the traditional *HST* flux calibration methodology. According to Re05, the *Spitzer* flux calibration is defined such that a point-source flux estimate

$$\langle F_\nu \rangle = C' N_e = F_{\nu_o} K, \tag{1}$$

where $\langle F_\nu \rangle$ is the mean flux over the bandpass and C' is the calibration constant for a point source. N_e is the number of detected photo-electrons per second, either N_e(pred) predicted from the stellar flux and the system fractional throughput R or N_e(obs) observed in an infinite-radius photometric aperture. Re05 uses a 10 pixel (12″) reference radius for the published calibration constants; but these calibrations refer to surface brightness (see Appendix A.2 and Equation (A10)). F_{ν_o} is the flux at the nominal wavelength $\lambda_o = c/\nu_o$ for a $\nu F_\nu = constant$ flux spectrum. N_e(pred) is

$$N_e(\text{pred}) = A \int \frac{F_\nu}{h\nu} R \, d\nu = \frac{A}{hc} \int F_\lambda \lambda R \, d\lambda, \tag{2}$$

where $A = 4869$ cm^{-2} is the collecting area of the *Spitzer* 85 cm primary mirror with its 14.2% obscuration (Werner et al. 2004).

K is the color correction

$$K = \frac{\int (F_\nu / F_{\nu_o})(\nu/\nu_o)^{-1} R \, d\nu}{\int (\nu/\nu_o)^{-2} R \, d\nu}, \tag{3}$$

where F_ν is the actual stellar spectral flux distribution. The nominal wavelength is

$$\lambda_o = \frac{\int \lambda \nu^{-1} R \, d\nu}{\int \nu^{-1} R \, d\nu} = \frac{\int R \, d\lambda}{\int \lambda^{-1} R \, d\lambda}, \tag{4}$$

where the term after the first equal sign is from Re05 and the term after the second equal sign is the equivalent formulation of Hora et al. (2008).

If the SEDs of the *HST* stars in Table 1 are used for the flux F_ν to produce a new calibration constant C'_{ST} and corresponding mean flux $\langle F_\nu{}^{ST} \rangle$, then the ratio of the new to the original calibration is

$$\frac{\langle F_\nu^{ST} \rangle}{\langle F_\nu \rangle} = \frac{C'_{ST}}{C'} = \frac{\int F_\nu / \nu R \, d\nu}{\nu_o \langle F_\nu \rangle \int \nu^{-2} R \, d\nu} \tag{5}$$

or, equivalently, in terms of integrals over wavelength

$$\frac{C'_{ST}}{C'} = \frac{\lambda_o \int F_\lambda \lambda R \, d\lambda}{c \langle F_\nu \rangle \int R \, d\lambda}, \tag{6}$$

where $\langle F_\nu \rangle$ is the stellar flux derived with aperture photometry from the *Spitzer* images, as calibrated with the official calibration constants that appear in the data-file headers. Re05 quotes an uncertainty of 2% in the IRAC absolute flux calibration. Tests of the numerical integrations over ν as per Equation (5) or over λ as per Equation (6) are the same to a few parts in 10^5.

2.2. Simplified Concept of a Point-source Calibration

A specific-intensity calibration for the surface brightness (see Appendix A.2) of diffuse sources requires a measure of the total response to a point source in an infinite aperture, as specified above for N_e. However, a flux calibration that is strictly for point sources has no such requirement and can be explained simply and elegantly in the case of a stable instrumental configuration with a linear response. Stability means that repeated observations produce the same response, measured in terms of say a background-subtracted net count rate N, while linearity implies that the count rate is directly proportional to the physical flux F, i.e., the ratio of flux to count rate will be the same ratio of F/N over the dynamic range of the system. There is no restriction on the entrance slit or extraction aperture as long as the same choice is made for both stars and the extracted count rate is repeatable for both stars. The measured count rate can be in a certain radius aperture for point-source photometry or of a certain height for a resolution element of a spectrophotometer. If one star is a flux standard with known flux, then the ratio F/N defines a point-source calibration constant P, so that the second star with unknown flux has the same constant measured ratio and the unknown flux is simply $F = PN$. This constant P might be alluded to as a sensitivity but is really more properly an inverse sensitivity, because a more sensitive instrument will have a higher count rate for a source of the same flux.

The main complication of this concept is due to the different spectral resolutions of the flux standard and the unknown star. A common example is a standard star with a tabulated medium resolution SED. For broadband photometry, the average flux of the standard over the bandpass must be calculated as in Equation (A1) or (A2), which is straightforward if the spectral resolution is much better than the bandwidth. In the case of a spectrometer calibration with a resolution that is lower than the

Table 2
Spitzer Calibration Parameters

Channel	λ_o (μm)	Cal. Ap. (pixels)	Ap. Corr.	Ap. Ref.[a]	R Ref.[b]
IRAC1	3.544	10	1.112	1	3
IRAC2	4.487	10	1.113	1	3
IRAC3	5.710	10	1.125	1	3
IRAC4	7.841	10	1.218	1	3
IRSB	15.793	∞	1.56	2	4
MIPS24	23.675	∞	...[c]	...	5

Notes.
[a] References for aperture corrections: (1) Hora et al. 2008; (2) IRS Instrument Handbook, 4.2.3.1 (http://ssc.spitzer.caltech.edu/irs/irsinstrumenthandbook/45/).
[b] References for the system-throughput spectral-response curves, R: (3) http://ssc.spitzer.caltech.edu/irac/calibrationfiles/spectralresponse/; (4) http://ssc.spitzer.caltech.edu/files/spitzer/bluePUtrans.txt divided by 3.58; (5) http://ssc.spitzer.caltech.edu/files/spitzer/MIPSfiltsumm.txt.
[c] PSF photometry is used for MIPS.

tabulated resolution of the standard, the calibration P is defined simply as the known SED, F, binned to the bandpass of the instrument to be calibrated divided by the response spectrum N for the same standard star. More properly, P as a function of wavelength is defined as the convolution of the known SED, F, with the instrumental line-spread function (LSF) divided by the count rate spectrum, N, of the standard convolved with the LSF of the standard star spectrum, which brings the numerator and denominator spectra of P to the same resolution and enables a pixel-by-pixel division of F by N after resampling to the same wavelength scale. This procedure may fail for the case where a low resolution standard star SED must be bootstrapped to a calibration of a much higher resolution spectrometer where the sensitivity of the high resolution data changes significantly over the resolution element of the known SED. For example, a single echelle order may have a variation in sensitivity by a factor of 10 or more over a wavelength range covered by only one or a few resolution elements of the flux standard.

3. *SPITZER* CALIBRATION

Spitzer observations of the sources were taken either as part of a cycle 5 Director's Discretionary Time program (PI: K. Gordon) or from existing archival observations. The stars were observed in the four IRAC bands, and as many as possible were observed in the IRS blue peakup band or MIPS 24 μm band. The main goal of the IRS blue peakup and MIPS 24 μm observations is to check for dust emission (e.g., a debris disk) or faint red companions.

3.1. Data Reduction

All the IRAC and IRS blue peakup data (reduction version S18.7.0) were downloaded from the *Spitzer* archive. The photometry is measured using a 3 pixel radius aperture and sky annulus with radii of 10 and 20 pixels on each individual image. As our sources are faint, refined positions are determined by centroiding on the star in each observation mosaic image. An aperture of three pixels radius is chosen in order to minimize noise from the sky and contamination from other sources in our relatively crowded fields. To get the total stellar flux, $\langle F_\nu \rangle$, our 3 pixel radius aperture photometry is corrected to the standard calibration aperture sizes as per Table 2.

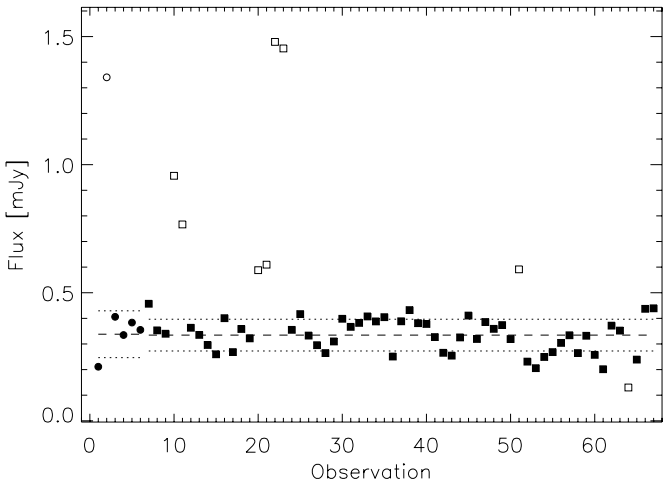

Figure 1. IRAC4 photometry for G191B2B for the first (circles) and second (squares) AOR. The filled symbols gives those measurements that were used in computing the averages (dashed lines) and standard deviations (dotted lines). The open symbols are the measurements that were iteratively sigma-clipped because of contamination by cosmic ray hits.

Table 2 includes the nominal wavelengths computed from Equation (4), the size of the standard reference aperture for each filter (i.e., 10 pixels for IRAC and infinite for IRSB and MIPS), the aperture correction needed to convert our three pixel photometry to the reference aperture size, and the references for the aperture corrections and instrumental-throughput spectral-response R versus vacuum wavelengths. Our aperture correction values for the IRAC photometry of HD 165459 agree with the tabulated values from Hora et al. (2008) to better than 0.25%, even though the Hora background annulus is 10–20 pixels instead of the 12–20 pixels used by Re05 for the standard 10 pixel photometry. The IRAC aperture correction from 10 pixels to infinity is discussed in Appendix A.2. Our nominal vacuum wavelengths λ_o are within 0.4% of those in Re05 and within 0.2% of Hora et al. (2008).

For G191B2B, a nearby bright star produces an artifact in its sky region; and the pixels affected are rejected prior to determining the sky flux. For each independent observation (Astronomical Observation Request, AOR), the photometry from the multiple image frames is averaged after sigma-clipping rejection of outlying points. For example, Figure 1 illustrates the rejected and accepted IRAC4 photometry points for the two AORs for G191B2B.

The weighted average IRAC and IRS fluxes for each star and band are given in Table 3 after multiplication of our three pixel radius photometry by the aperture corrections in Table 2. In order to achieve robust results, IRAC observations with a signal-to-noise ratio of less than 7 or with a location more than 25 pixels from the center are rejected and do not contribute to the averages in Table 3. The restriction to centrally located sources avoids any confusion due to possible errors in the flat fielding procedure. The second line for each star in Table 3 is the synthetic photometry predictions computed from our standard star SEDs as per Equation (A2).

3.1.1. Uncertainties

Many (~100) AORs exist for the four primary Re05 standard stars. For example, Figure 2 shows the observations of two of our stars, HD 165459 and 1812095, that are also Re05 primary standards. Each point in Figure 2 represents the sigma-clipped average of multiple image frames in one AOR. The

Table 3
Spitzer and Synthetic Photometry[a] (mJy)

Star	IRAC1		IRAC2		IRAC3		IRAC4		IRSB		MIPS24	
	Flux	Unc(%)	Flux	Unc(%)	Flux	Unc(%)	Flux	Unc(%)	Flux	Unc(%)	Flux	Unc(%)
G191B2B	2.04e+00	0.64	1.29e+00	0.61	8.05e-01	1.24	4.08e-01	2.60	...	...	...	...
	2.06e+00		1.31e+00		8.25e-01		4.57e-01		...		...	
GD153	5.01e-01	0.70	3.12e-01	0.84	1.95e-01	4.16	1.07e-01	7.11	...	...	...	...
	4.96e-01		3.13e-01		1.96e-01		1.08e-01		...		...	
GD71	6.93e-01	0.69	4.39e-01	0.67	2.79e-01	3.26	1.45e-01	4.72	...	...	...	...
	6.76e-01		4.25e-01		2.66e-01		1.45e-01		...		...	
LDS749B	2.52e-01	0.84	1.59e-01	1.24	1.11e-01	6.08	5.56e-02	12.34	...	...	...	...
	2.65e-01		1.70e-01		1.07e-01		5.93e-02		...		...	
HD 165459	6.62e+02	0.08	4.26e+02	0.09	2.69e+02	0.17	1.48e+02	0.12	...	...	2.57e+01	0.29
	6.50e+02		4.21e+02		2.69e+02		1.50e+02		...		1.70e+01	
1732526	3.65e+00	1.16	2.34e+00	1.74	1.50e+00	3.70	...	...	1.58e-01	6.03	...	...
	3.58e+00		2.32e+00		1.48e+00		...		2.10e-01		...	
1740346	4.62e+00	0.86	3.02e+00	1.13	1.97e+00	1.78	1.12e+00	1.00	6.59e-01	2.93	...	...
	4.45e+00		2.89e+00		1.85e+00		1.03e+00		2.64e-01		...	
1743045	2.13e+00	0.22	1.38e+00	0.27	8.88e-01	0.43	4.83e-01	0.41	...	...	...	...
	2.07e+00		1.34e+00		8.59e-01		4.77e-01		...		...	
1802271	5.19e+00	0.26	3.35e+00	0.31	2.16e+00	0.58	1.16e+00	0.36	2.26e-01	6.55	...	...
	5.12e+00		3.32e+00		2.12e+00		1.17e+00		3.00e-01		...	
1805292	4.58e+00	0.95	2.99e+00	1.09	1.93e+00	3.55	1.07e+00	1.07	2.28e-01	4.61	...	...
	4.43e+00		2.88e+00		1.84e+00		1.02e+00		2.62e-01		...	
1812095	8.85e+00	0.07	5.75e+00	0.09	3.66e+00	0.19	1.98e+00	0.16	4.96e-01	4.68	2.04e-01	8.82
	8.61e+00		5.59e+00		3.58e+00		1.99e+00		5.10e-01		2.26e-01	
HD 209458[b]	...	...	...	...	...	...	...	...	...	...	2.13e+01	0.11
	...		...		...		...		...		2.21e+01	
P041C	1.79e+01	0.77	1.12e+01	0.97	7.39e+00	1.56	4.12e+00	2.07	1.14e+00	6.64	...	...
	1.73e+01		1.10e+01		7.09e+00		3.98e+00		1.02e+00		...	
P330E	7.84e+00	0.73	5.07e+00	0.88	3.21e+00	3.43	1.80e+00	1.65	...	...	2.20e-01	5.45
	7.78e+00		4.97e+00		3.20e+00		1.79e+00		...		2.04e-01	

Notes.
[a] The first line for each star contains the measured fluxes, while the second line is the predicted synthetic photometry.
[b] Existing IRAC and IRSB data are only in non-standard modes.

rms indicated on each panel is the scatter among the remaining AOR observations after rejecting points that deviate by more than 3σ from the average. The brighter star HD 165459 has 2, 5, 4, and 1 rejected deviant AORs with an rms for the remaining points of 0.6%, 0.7%, 0.6%, and 0.8% for channels 1–4, respectively, while the comparable rms dispersions from Re05 are 1.7%, 0.9%, 0.9%, and 0.5%. In order to compute the weight for each AOR included in the average fluxes of Table 3, our repeatabilities for HD 165459 are added in quadrature with the statistical uncertainty for each independent observation. Figure 3 compares the *Spitzer* broadband fluxes to our absolute flux distributions for one example of each of our three stellar classes. The differences between the flux levels for the nominal and effective wavelengths quantify the ambiguity associated with assigning monochromatic wavelengths to the broadband *Spitzer* photometry.

3.1.2. Comparison with the Re05 Photometry

Re05 based the final, recommended IRAC calibration on four primary A star standards, two of which, HD 165459 and 1812095, are among our standard stars. Table 1 of Re05 contains the photometric fluxes for these two stars along with their other two primaries, HD 180609 and BD+60°1753. On average, our extracted IRAC photometric fluxes are 2.5%, 2.8%, 2.4%, and 1.1% higher than the corresponding Re05 tabulations for channels 1–4, respectively.

Our IRAC photometry from each image is corrected for distortions and pixel phase as recommended by Hora et al. (2008). These corrections differ from Re05, who used a preliminary version of the Hora et al. work. Because of these different data reduction procedures, our corrected photometry is expected to be systematically brighter than Re05 by 1.1%, 1.4%, 0.6%, and 0.5%, for bands 1–4, respectively. These expected systematic difference between Re05 and our photometry account for around half of the actual differences. The remaining unexplained differences of up to 1.8% for IRAC3 must be due to the changing IRAC pipeline processing and/or a somewhat different selection of IRAC observations to include in the average for each star.

3.1.3. Special Cases

There are no IRAC4 measurements for 1732526. Two additional *HST* standard G stars have IRAC observations (C26202 and SNAP-2), yet neither star has high enough quality observations to be included in this work. C26202 is in the CDF-S (Smith et al. 2003) and has a cooler companion at 3″–4″ that produces blended IRAC images and precludes accurate photometry of the separate stars. The second G star, SNAP-2, has IRAC data but lies off-center by ∼50 pixels where flat fielding errors might be important.

3.1.4. MIPS

The raw MIPS data were downloaded from the *Spitzer* archive and reduced using the MIPS Data Reduction Tool (Gordon et al. 2005). In addition, several additional steps to remove residual instrumental signatures are used (see Engelbracht et al. 2007 for details). Given the crowded nature of some of the fields,

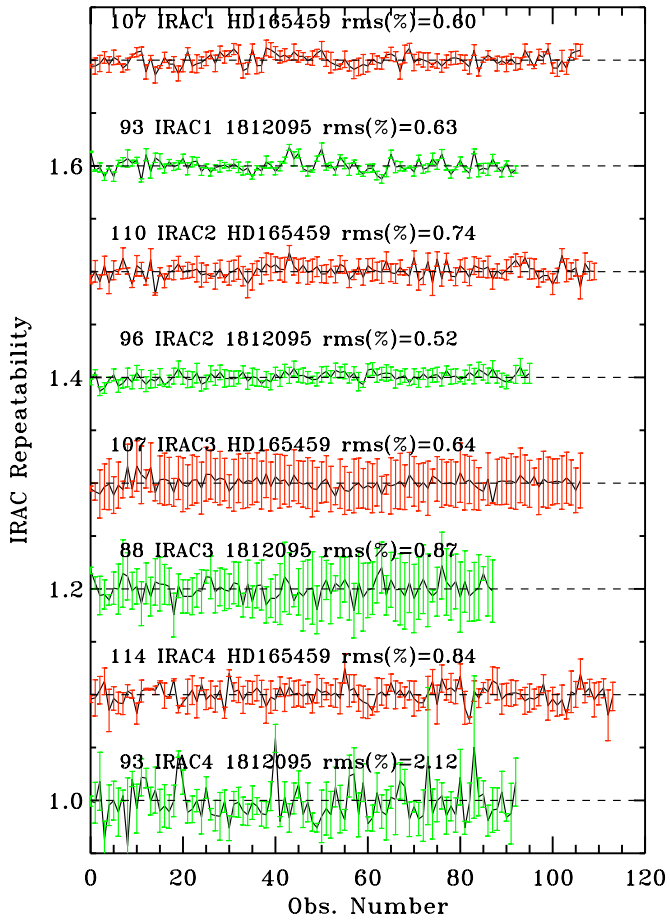

Figure 2. Observations of two Re05 primary stars in the four IRAC bands, where each point is the result from one AOR. Each set of points is offset by 0.1 along the *Y*-axis from the set below. The points displayed all have an S/N of at least 7 and are within 25 pixels of image center in each IRAC channel. The red points are for HD 165459, while 1812095 is shown in green. Each of the eight panels is labeled with the number of observations (AORs), the IRAC band, the star name, and the rms of the points shown.

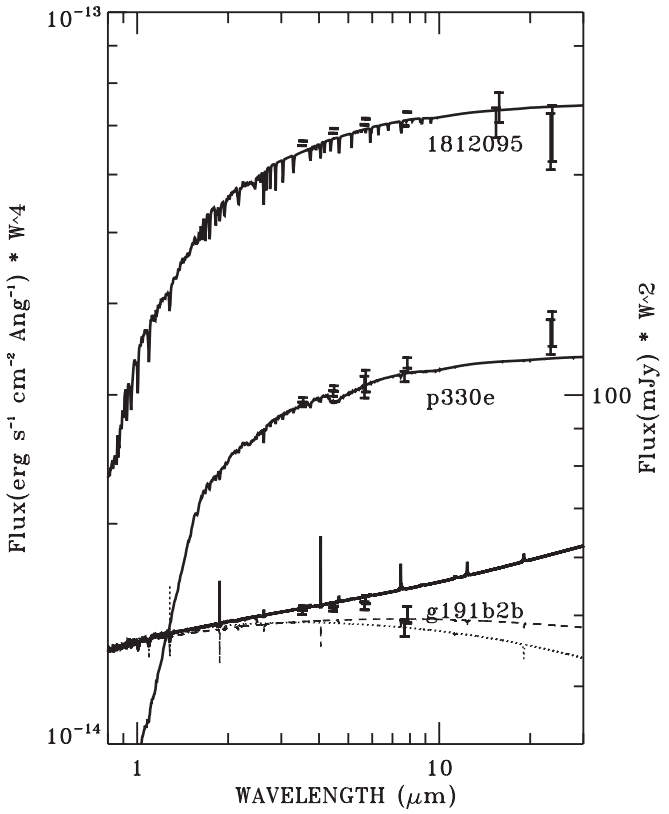

Figure 3. Continuous lines are SEDs typical of our three spectral categories of standard stars scaled by λ^4 on the left axis and by λ^2 on the right axis with λ in μm. The absolute fluxes are multiplied by a factor of two for G191B2B and 1812095. The statistical error bars of $\pm 1\sigma$ are shown for the measured Spitzer photometry at the nominal wavelengths and at the effective wavelengths. For G191B2B, a Tlusty LTE model (dash) and a NLTE 60,000 K model (dots) with solar CNO (Gianninas et al. 2010) are shown in addition to the standard NLTE model (solid).

point-spread function fitting code is required to extract the MIPS24 photometry of our sources; and our choice is StarFinder (Diolaiti et al. 2000).

3.1.5. Predicted Fluxes

The measured and synthetic fluxes appear in alternate rows in Table 3. For the three pure hydrogen WDs, the temperature and gravity of the TLusty NLTE models are derived from fits to ground-based spectra of the Balmer lines. For the pure He WD LDS749B, the model of Bohlin & Koester (2008) is used. For the A stars, Bohlin & Cohen (2008) fit NICMOS spectrophotometry from 0.8 to 2.5 μm and ground-based photometry with Castelli & Kurucz (2004, hereafter CK04) model SEDs. Similarly for the G stars, Bohlin (2010, hereafter B10) fit STIS and NICMOS spectrophotometry. Comparing our Table 3 synthetic fluxes with the corresponding values in Table 6 of Re05 for the two stars in common confirms that Re05 used stellar SEDs without modification from the Cohen CWW network. A generous statistical uncertainty of 1% is assigned to our Table 3 synthetic fluxes to account for the effects of the broadband Poisson noise and repeatability of STIS spectrophotometry (Bohlin 2002) on the fitting of models to the *HST* spectrophotometry. More details of the *HST* synthetic fluxes in Table 3 and their systematic uncertainties are discussed in Section 4.1.

3.2. Linearity

The linearity of five of the *Spitzer* detector systems can be evaluated by comparing the observed photometry to the actual stellar brightness. A more common linearity check is a comparison of count rate versus well depth. For the sixth system, MIPS24, our three data points are insufficient to reach any conclusion; see Engelbracht et al. (2007) for details of the confirmation of MIPS 24 μm linearity. Using the modeled *HST* SEDs as the stellar flux, Equation (5) is the ratio of predicted to measured fluxes; and that ratio is shown in Figure 4 as a function of stellar brightness for our three classes of stars. IRAC3 has the narrowest bandpass and the lowest count rates. G191B2B is significantly high in the IRAC4 band; but that one anomalous point is not an indication of non-linearity. However, for IRSB, even over the small dynamic range of 7, the five data points show evidence of some issue with the photometry. The ratio for the faintest IRSB source 1732526 is 30% high, while the brightest source P041C is 10% low. However, Gilliland & Rajan (2011) discovered that P041C is a double star with an M star companion separated by 0.″57, which could make the observed IRSB flux too high by ∼10%. These problems preclude an accurate comparison of the *HST* versus *Spitzer* IRSB calibration.

3.3. Dust Rings

Many stars are encircled by a ring of cool dust that emits strongly at longer wavelengths. These sources can be identified by small ratios of predicted to observed flux that fall off-scale

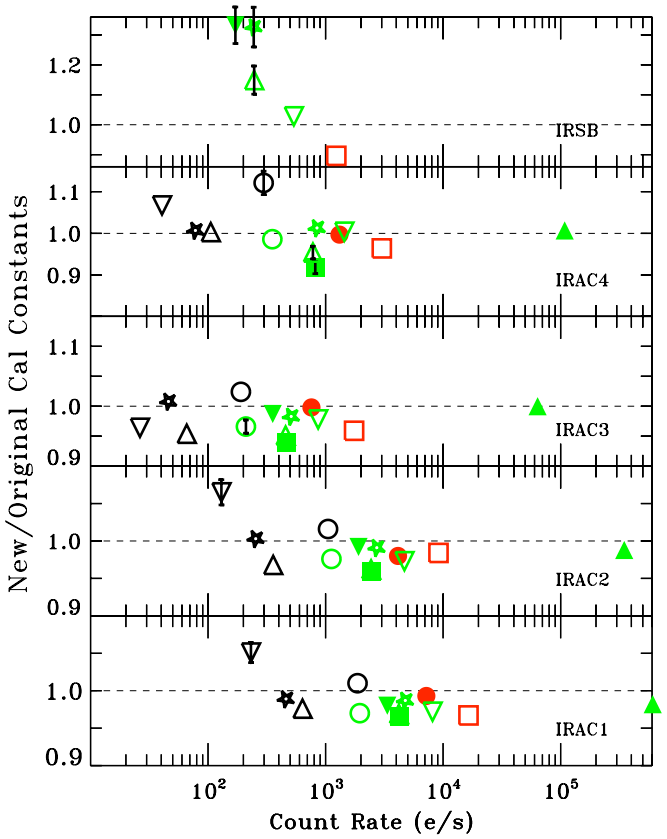

Figure 4. Check of the linearity of the observed IRAC and IRSB photometry versus source brightness. The plotted ratios per the equivalent Equations (5) and (6) are a measure of the mean flux of the *HST* SEDs divided by the measured photometric flux per the *Spitzer* pipeline calibrations. The stars are color coded by spectral type, with black, green, and red for WDs, A stars, and G stars, respectively. One sigma error bars are shown on the points that differ from unity by more than 3σ. The key for each symbol type is at the right side of Figure 5. Note the progressively more compressed scales in the upper three panels.

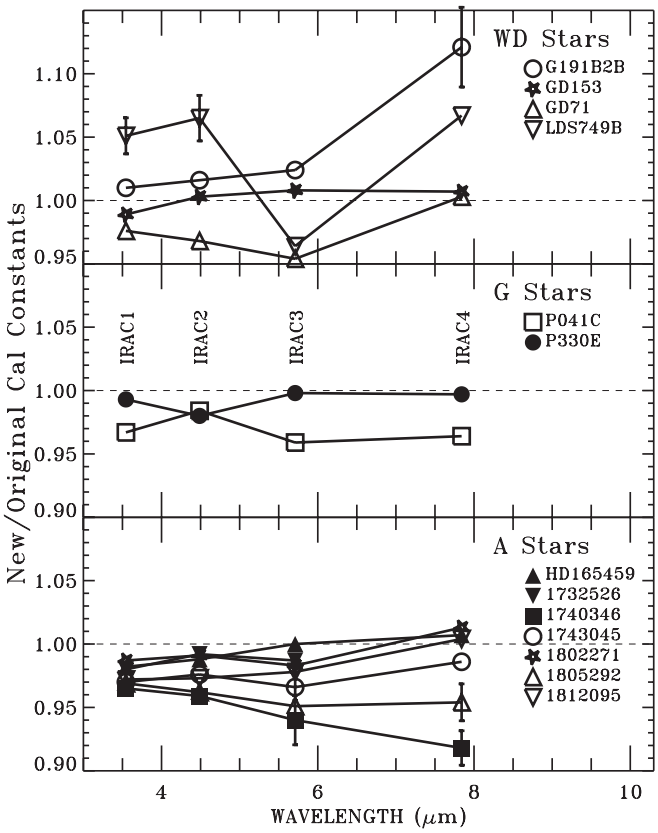

Figure 5. Ratios as in Figure 4 of calibration constants computed from the *HST*-based fluxes to the IRAC values of Re05. WDs are in the top panel, solar analogs are in the center panel, and A stars are at the bottom. One sigma error bars are shown on the points that differ from unity by more than 3σ. The key to the individual symbols is at the right side of each panel.The scales are the same for all three stellar types, but there is an offset of 0.05 in the top panel.

in Figure 4. Our observed excess flux at 24 μm for HD 165459 confirms the evidence for a disk found by Rieke et al. (2005), who reported a 40% excess. Su et al. (2006) revised the excess to 46.7%.

Our excess of a factor of 2.5 for 1740346 in the 16 μm band suggests the presence of a contaminating debris disk. While there is no evidence of dust emission below 8 μm for HD 165459, the low values for 1740346 in Figure 5 suggest the presence of hotter dust than is around HD 165459. Therefore, 1740346 is not included below in the comparison with the Re05 calibration of IRAC.

3.4. Results

As per Equation (5) or (6), Figure 5 compares the *HST* calibration constants to those published by Re05. One σ error bars appear on the points that differ from unity by more than the 3σ uncertainty of the ratio. The problematic IRSB data are not shown; and the MIPS24 ratios for the three stars without dust rings are insufficient to draw any firm conclusions, although all three MIPS ratios are within ~10% of unity.

With a 4σ statistical significance, the IRAC4 ratio of 1.12 for the observation of G191B2B is the largest deviation from unity in Figure 5. As a check on this problematic case, the background level in the IRAC4 band was studied by constructing a histogram of the set of all three-pixel radius photometry in a 120 pixel square, centered on G191B2B. There are no background regions in the vicinity of the star that are low enough to bring the observed result to the predicted level with a ratio of unity in Figure 4; and there is only ~1% probability of a sky level that would reduce the discrepancy to 2σ.

Table 4
C'_{ST}/C' = New/Re05 IRAC Calibration Constants

Channel	Star Set	New/Orig.	Unc.
IRAC1	4 WD	1.004	0.006
IRAC1	2 Primary	0.977	0.007
IRAC1	8 A+G	0.977	0.004
IRAC2	4 WD	1.006	0.006
IRAC2	2 Primary	0.981	0.007
IRAC2	8 A+G	0.981	0.004
IRAC3	4 WD	1.007	0.013
IRAC3	2 Primary	0.988	0.007
IRAC3	8 A+G	0.980	0.005
IRAC4	4 WD	1.077	0.024
IRAC4	2 Primary	1.005	0.007
IRAC4	7 A+G	0.995	0.005

Table 4 reports the average results separately for the WDs, for two Re05 primary A stars, and for the group of A and G stars. These values are the weighted averages of the ratios of the synthetic to the measure fluxes in Table 3, where the uncertainty for each star is the measurement error in Table 3 combined with the 1% uncertainty assigned to the synthetic values. These uncertainties are the statistical errors and do not include any possible systematic errors that would affect all stars equally. As a group, the averages for the WDs are significantly above the

other averages, especially with the inclusion of the IRAC4 ratio for the hottest WD, G191B2B, that is 12% high in Figure 5. Individually, many of the WD results could be explained as not significant at the 3σ level or because of the less well vetted pure helium model for LDS749B (Bohlin & Koester 2008).

NLTE effects in the IR become more pronounced at higher stellar temperatures (Bohlin 2000), e.g., the pure hydrogen LTE model for G191B2B shown in Figure 3 falls near the IRAC4 data. For G191B2B, metal line-blanketing is observed at the 1%–2% level in the FUV (Barstow et al. 1999). As per Gianninas et al. (2010), an additional NLTE model spectrum at 60,000 K and $\log g = 7.5$ with CNO metals at the Asplund et al. (2005) solar abundance also appears in Figure 3 near the LTE model. This metal abundance is not the actual metallicity but is only a proxy for the effect of the metals, which is to reduce the NLTE effects and cool the upper atmosphere, where the IR continuum is formed. There is no coherent simultaneous determination of the metal abundances along with $T_{\rm eff}$ and $\log g$ from the Balmer lines (Barstow et al. 2003); however, the abundances suggested by Barstow et al. (2003) are considerably less than the solar Asplund et al. (2005) values, so that a model with proper trace metallicities for G191B2B in the IR should fall somewhere between the 60,000 K CNO model and the pure hydrogen NLTE model, potentially within $1–2\sigma$ of all the IRAC fluxes.

The significant deviation from unity in Figures 4–5 for G191B2B should be explored with a proper model. Meanwhile, the most relevant result in Table 4 is for the set of eight G and A stars, for which our calibration constants are lower than Re05 for IRAC1–4 by 2.3%, 1.9%, 2.0%, and 0.5%, respectively. For the four Re05 primary stars, our re-reduced photometry is 2.5%, 2.8%, 2.4%, and 1.1% higher than Re05. Higher extracted photometry implies lower calibration constants; and the measured photometry differences correspond to the calibration constant differences to an accuracy of better than 1%. Understanding that our differences with Re05 are mostly due to differences in the photometry extracted from the IRAC images helps verify that both calibrations have been properly derived per the adopted common methodology and that our results are expected to be lower by about the amounts computed in Table 4 for A+G stars.

While the calibration constants themselves depend directly on the accuracy of the bandpass throughput R as per Equation (A1) or (A2), our methodology and that of Re05 involve the same integral of the stellar SED over R, so that errors in R cancel to first order in the comparison of the two sets of results. However, bandpass errors, such as an overall shift of R in wavelength, change the predicted N_e(pred) as per Equation (2). For example, a shift of $+0.1\,\mu$m in the $2.9\,\mu$m wide IRAC4 band, i.e., a shift in wavelength by 3% of the bandwidth, would cause a 3.8% decrease in N_e(pred) for the WD, A, and G stellar types. As per Equation (2), such a 3.8% error would be difficult to distinguish from a simple 3.8% compensating error in the laboratory measurements of R, because all of our standard stars have nearly the same Rayleigh–Jeans slope in the IR. A bandpass shift would only be important in the case of strong spectral features within the bandpass or in the case of a source with an SED much cooler than our A and G standard stars.

4. DETAILS OF ABSOLUTE FLUX CALIBRATION

4.1. Details of the HST Calibration

The discussion of our measures of instrumental response (N) appears in Sections 2.1–2.2, while this section covers the details of the flux (F) in the calibration constant $P = F/N$.

The essence of the *HST* flux system is to establish standard star SEDs and then measure other stars relative to these standard candles. Pure hydrogen WD stars are chosen as these *HST* fundamental standard candles, because their atmospheric models are simpler than other stars where metal lines, molecular lines, and convection add complications. The models of our pure-hydrogen primary WD standards are specified by two parameters $T_{\rm eff}$ and $\log g$, both of which are defined by the Balmer line profiles. To establish absolute flux standards, the unreddened WD stars G191B2B, GD71, and GD153 are observed with STIS, which also observed Vega in the same modes with the same dispersion. Because STIS is precisely linear even into the regime of many times over-saturated in its CCD (Gilliland et al. 1999; Bohlin & Gilliland 2004), the relative flux between Vega and each WD is measured (Bohlin & Gilliland 2004); and the well-known absolute monochromatic flux at 5556 Å for Vega establishes the absolute flux of the model SED for each WD. Minor complications arise due to the slowly changing STIS sensitivity with time and the gradual loss of charge transfer efficiency in the CCD. These corrections are tracked to better than 1%, and the STIS response is corrected for these effects as a function of wavelength.

Observations of the three fundamental primary WDs establish the instrumental flux calibration. For example, a set of observations in the low dispersion modes of STIS below $1\,\mu$m and from 0.8 to $2.5\,\mu$m in the grism modes of NICMOS establish the flux calibrations and enables the creation of secondary flux standards. Bohlin & Koester (2008) demonstrated that such STIS and NICMOS spectrophotometry of LDS 749B, a pure helium star, could be modeled to the statistical precision of the data. NICMOS spectrophotometry, supplemented by ground-based photometry for the A stars used in this paper, was modeled by Bohlin & Cohen (2008) with $T_{\rm eff}$, $\log g$, $\log z$, and color excess $E(B-V)$ as free parameters. Similarly, B10 modeled STIS and NICMOS fluxes for the G stars discussed here. The *BVR* photometry used to define the A star models is from Mount Hopkins Observatory (Cohen et al. 2003) and is referenced to the 9400 K Vega model and to Vega photometric values from Maíz Apellániz (2007).

4.1.1. Uncertainties

Establishing proper uncertainties is a complex process, involving the division into categories of systematic and statistical errors, which can each be divided into sub-categories of absolute level and relative slope of flux versus wavelength (i.e., "color"). In many cases, the statistical scatter can be reduced below the systematic error bars by repeated observations or by binning spectra. Systematic uncertainty of the *HST* WD flux scale is estimated as ~1% by B10 for the ratio of the flux at 5556 Å to fluxes in the 1–2.5 μm range. The total systematic uncertainty in our IR flux scale should be less than 2%, even when the 0.7% uncertainty in the absolute 5556 Å flux (Megessier 1995) is included. Considerable confidence in the estimate of precision in the slope of the relative flux with wavelength can be gained by examining the internal agreement among the *HST* standards: (1) Bohlin (2007a) illustrates the $\ll$1% agreement of the three primary WD models with their calibrated fluxes, which means that if the slope of any model, which depends primarily on $T_{\rm eff}$, is in error, then the other two model temperatures must be similarly in error in order to make the same change in flux versus wavelength for all three stars from 0.12 to $2.5\,\mu$m. (2) The NICMOS photometric calibration is based on the *HST* primary standard G191B2B and the secondary

standard P330E as per de Jong (2006), where the consistency of the absolute spectrophotometry is confirmed for these two SEDs from the CALSPEC[8] database. In particular, de Jong's Figure 2 shows an agreement of 0.8%–1.4% between the NICMOS calibrations derived from the SEDs for P330E and G191B2B over the combined 0.9–2.4 μm range of the three cameras. Finally, the extrapolation of the *HST* NICMOS fluxes from 2.2 μm to the first IRAC band at 3.6 μm is rather independent of the particular stellar model. For example, among the seven A stars in Table 1, Bohlin & Cohen (2008) demonstrate that all the measured NICMOS SEDs fit their model SEDs over the 0.8–2.4 μm range within 1% in broad wavelength bins; and the maximum difference in the 2.2/3.6 μm ratio is 0.7% between the minimum $T_{\rm eff}$ = 7650 K model for 1743045 and the maximum 9100 K model for 1802271.

4.2. Details of the Re05 Calibration

As discussed above, our measured IRAC instrumental responses are smaller than published by Re05 for the same stars. However, Re05 used data from the IRAC pipeline processing version S10, while version S18.7.0 is utilized for this paper. To be directly comparable, any two independent calibrations should use the same pipeline processing version, as well as the same algorithms for extracting the point-source fluxes from the images.

As discussed above and in the Appendix, the equations that define our new calibrations are exactly equivalent to those of Re05, so that differences arise only from different measured instrumental photometry or from differences in the adopted stellar fluxes.

The stellar fluxes used by Re05 are from the CWW grid, while the revised SEDs from Bohlin & Cohen (2008) are used here. Some major differences in the derivation of model SEDs are that Bohlin & Cohen fitted models to NICMOS spectrophotometry in the 0.8–2.4 μm range rather than to the Two Micron All Sky Survey plus *I*-band photometry used for determining the CWW SEDs. The CK04 model grid was used to fit the data; and rather than specifying the model $T_{\rm eff}$ and log g from spectral classification spectra, Bohlin & Cohen found the best fitting model from the grid, allowing $T_{\rm eff}$, log g, log z, and the reddening $E(B-V)$ to vary as free parameters. Despite these differences, the resulting SEDs differ only slightly from the CCW fluxes used by Re05. For example, over the IRAC 3–9 μm range, our SEDs for the IRAC primary stars HD 165469 and 1812095 agree with the CWW SEDs used by Re05 to 0.5 and 1%–2%, respectively.

4.3. Details of the Ri08 Calibration

Ri08 start by determining a best calibrated value for an equivalent Vega photospheric flux at 10.6 μm and use a Vega theoretical model normalized to this value to predict the photospheric flux density at 2.22 μm. This prediction is robust against different models for A0 star photospheres. After correction for the small contribution for the extended debris ring found in interferometry, the resulting prediction is compared with measurements of the absolute flux density from Vega near 2.22 μm. The agreement is good; and the two approaches are averaged to generate a "best" A star-based calibration at 2.22 μm. Ri08 independently generated a calibrated solar spectrum as a combination of the measurements of Thuillier et al. (2003) out to 2.4 μm and an Engelke function from 2.4 to 12 μm (over which range the Engelke function gives a good fit to a number of accurate calibrated solar measurements). Ri08 use this spectrum to compare with the observed $K-[8]$ color of solar-type stars and extrapolated the result to 10.6 μm for an independent test of the beginning calibration at this wavelength. Their calibration makes no reference to the visible calibration, although the predicted color of Vega based on the absolute calibrations at *V* and *K* is consistent with the observed color of the star to within about 2%. The $V-K$ colors of solar-type stars are also consistent roughly within this error with their calibration at 2.22 μm using the Thuillier measurements to translate to *V*. The Ri08 results should be understood as a purely infrared-based calibration with an estimated accuracy in this region of 2%.

The Ri08 calibration suggests that the IRAC4 8 μm fluxes should be 1.5% $\pm$ 2% higher, while our fluxes are ~0.5% lower than Re05. Thus, our results agree with Ri08 within the uncertainties. A new extraction of the IRAC photometry for the 32 Ri08 stars compared with the Ri08 tabulation of predicted fluxes produces a mean predicted-to-observed ratio that is 2.1% $\pm$ 0.5% above unity, i.e., confirming the Ri08 difference of 1.5% within the uncertainties.

4.4. Comparison of Methodologies

The following points compare the Ri08 technique to the similar *HST* and CWW methods for establishing standard star SEDs.

1. Ri08 establish two independent IR absolute flux determinations for A stars (Vega) and for G stars (the Sun), for which agreement provides a powerful confirmation of the results. Both the *HST* and CWW absolute levels are traceable only to the visible range for Vega.
2. In the Ri08 method, a large number of flux standards are more easily and cheaply established from existing accurate photometric systems, so that statistical uncertainties can be reduced well below systematic uncertainties. The other methods are more observationally expensive, especially in the case of the *HST* flux measurements.
3. Because of the similarity of SEDs at long wavelengths, errors between widely separated IR wavelengths are minimal. Ri08 normalized the adopted SEDs to IR absolute flux measurements, so that errors in the IR are minimized.
4. The shape of stellar SEDs short of 1.5 μm becomes increasingly dependent on the stellar temperature and on the metallicity. Thus, the uncertainty of the flux calibration below 1.5 μm grows with decreasing wavelength making the Ri08 technique less useful for calibration in the short wavelength (0.6–1.5 μm) range of the *James Webb Space Telescope* (*JWST*).

5. RECOMMENDATIONS

Several deficiencies in our prototype set of flux standards must be addressed before *JWST* can be calibrated to our goal of 1% precision at 0.6–30 μm. Required improvements include (1) better stellar atmosphere grids and a better understanding of their uncertainties, (2) updated and expanded lists of potential standard stars in each of our three spectral type classes, and (3) resolution of the proper normalization of these models to measures of absolute stellar fluxes in the visible and IR.

5.1. Model Atmosphere Grids

Perhaps the best way to assess systematic errors within a set of similar type stars is to compare independent sets of models

[8] http://www.stsci.edu/hst/observatory/cdbs/calspec.html/

to the measured flux distributions, which B10 did for the G stars by fitting both MARCS models (Gustafsson et al. 2008) and those of CK04. In broad continuum bins, both sets of models agree with the *HST* fluxes to ∼0.5%. However, both of these sets have serious deficiencies. The MARCS grid does not extend past 20 μm, while the CK04 grid has rather coarse wavelength spacing with a total of 1221 points and only one point between 10 and 40 μm. A third independent model grid, an update of the CK04 models with good wavelength resolution, and an extension of the MARCS grid beyond its current 20 μm limit would be ideal. Shortward of 20 μm, the best-fit CK04 models agree with the best-fit MARCS models to 1% for the G stars in broad continuum bands. Combining this 1% with a 1% systematic uncertainty in the *HST* WD flux scale results in the B10 estimate of a possible systematic 2% uncertainty of the broadband IR fluxes of individual G stars with respect to the *V* band.

For the A stars, a second independent model grid is needed, because the MARCS models are limited to maximum temperatures of 8000 K. Both the original CWW SEDs and the newer Bohlin & Cohen (2008) SEDs are extrapolations of the measured fluxes into the IR using models that are based on computer code traceable to R. Kurucz. The best check on these IR fluxes of the A stars was from the two models provided by T. Lanz at 9400 K and 8020 K. While the agreement of the Lanz SED with CK04 is within 2% for the 9400 K model, the Lanz SED is brighter than CK04 by 4% at 10 μm for the 8020 K model, which raises the question of the accuracy of model SEDs, in general. Because of the scatter among the WDs and because only two G stars are included, our results are based mainly on A stars. Furthermore, the final Re05 IRAC calibration is based entirely on A stars, so that more work on independent A star grids is essential.

5.2. Sample of Stars

Our goal of comparing results among different stellar types is compromised by the large scatter and offset of the average results for the WDs and by the poor significance of the G star result that is based on only two stars. Perhaps, warm-mission IRAC1 and IRAC2 observations of our four WD stars could illuminate the reality of their large scatter in the top panel of Figure 5. Among the A stars, Figure 5 shows the most scatter at 8 μm, where 1740346 is excluded because of evidence for dust at 16 μm. However, 1805292 has a 3σ deviation without a strong indication for dust. Thus, more A stars would reduce uncertainties, especially with respect to the pure A star calibration of Re05. The remaining two primary and primary-candidate A star standards of Re05 should be observed with STIS and modeled to achieve a more complete and precise comparison between the *HST* and Re05 A star results.

In order to confirm the A star comparison with IRAC, more G stars and more WD stars with good existing cold mission *Spitzer* observations are needed. Plenty of observations of brighter stars with both IRAC and MIPS 24 μm exist in the Spitzer archives, except for WDs. Thus, the WD category will be supplemented by late O or B type stars. This new set of standards is also needed to cover the large dynamic range and spectral variety required to calibrate the full suite of *JWST* instrumentation but will require new STIS spectrophotometry to provide the link to *HST*. ACS photometry, WFC3 photometry, and WFC3 grism spectrophotometry will also help in firmly establishing the tie to the *HST* flux scale, especially for the WFC3 data at wavelengths longward of the STIS cutoff at 1 μm.

5.3. Absolute Flux Level

While the absolute-flux zero-point of all *HST* standards is tied to the monochromatic value of flux for Vega of 3.46×10^{-9} erg cm^{-2} s^{-1} Å^{-1} at 5556 Å (Megessier 1995), there are valid absolute measures in the IR, as summarized by Ri08. Because Ri08 discuss and present a SED for Vega, the differences between our preferred Vega SED and the Ri08 SED are investigated as a possible explanation for the differences between the *HST* and Ri08 flux calibrations for IRAC. The SED of Vega has been measured by STIS at 0.17–1 μm by Bohlin & Gilliland (2004), who suggested that the Kurucz $T_{\rm eff}$ = 9550 K model[9] fits the STIS observations, while Bohlin (2007a) discovered that a change in the STIS non-linearity correction made a Kurucz 9400 K model[4] fit the observed flux much better both below the Balmer jump and in the 0.7–0.8 μm region, as illustrated in Figure 6 with red for the 9550 K and green for the 9400 K models. These two models and the reference continuum level are all normalized to 3.46×10^{-9} erg cm^{-2} s^{-1} Å^{-1} at 5556 Å (Megessier 1995). The blue line is the same 9550 K model but is normalized to the Ri08 IR value of 645 Jy ± 2.3% at 2.22 μm. The Ri08 fluxes below 1 μm are not expected to match the actual stellar flux and are not shown. The division of all the SEDs by the same smooth theoretical model continuum in Figure 6 illustrates the differences among the various flux distributions and also shows where the (mostly hydrogen) line blanketing complicates the comparisons. The STIS measurements (black line in Figure 6) and the 9400 K model (green line) agree to ∼1% in the unblanketed regions from below the Balmer jump to ∼1 μm.

Fortunately, the discrepancy of 1.1% at 2.22 μm between the *HST*-based fluxes and Ri08 is almost negligible and could be reduced by using the 9400 K model with a weighted average of the visible and IR normalizations. Another complication is that Vega is a pole-on rapid rotator (Peterson et al. 2006) with temperature zones in the 7900–10150 K range (Aufdenberg et al. 2006). Even though the line blanketing in the Aufdenberg et al. multi-temperature model is occasionally stronger than measured, the continuum levels of his computed pole-on SED (J. P. Aufdenberg 2006, private communication) track the ratio of the 5556 Å to the 2.22 μm absolute fluxes to 0.4%, i.e., much better than any uncertainties in the absolute flux measurements in the visible or IR. In summary, there is no reason to question the accuracy of either the 5556 Å flux of Megessier (1995) or the 2.22 μm flux quoted by Ri08.

A robust and straightforward comparison of a model SED for Vega with the measured absolute visible and IR fluxes is not possible because of the IR emission from the dust ring and because Vega is a pole-on rapid rotator. As pointed out by Cohen et al. (1992a), Sirius is a much better primary IR standard because of its low rotation speed and lack of a contaminating dust ring. Thus, one big step forward in resolving any possible offset between visible and IR absolute fluxes is to observe Sirius with STIS and fit a model, in order to directly compare the model with IR absolute fluxes, e.g., from the Mid-course Space Experiment (Price et al. 2004), who have ∼1% measurement accuracy in six bands from 4.3 to 21.3 μm. STIS spectrophotometry of Sirius will be somewhat more saturated than for Vega; but the same techniques of Bohlin & Gilliland (2004) for precisely analyzing saturated data should apply.

On longer time scales, there are ongoing programs to measure stellar spectrophotometric fluxes relative to NIST laboratory

[9] http://kurucz.harvard.edu/

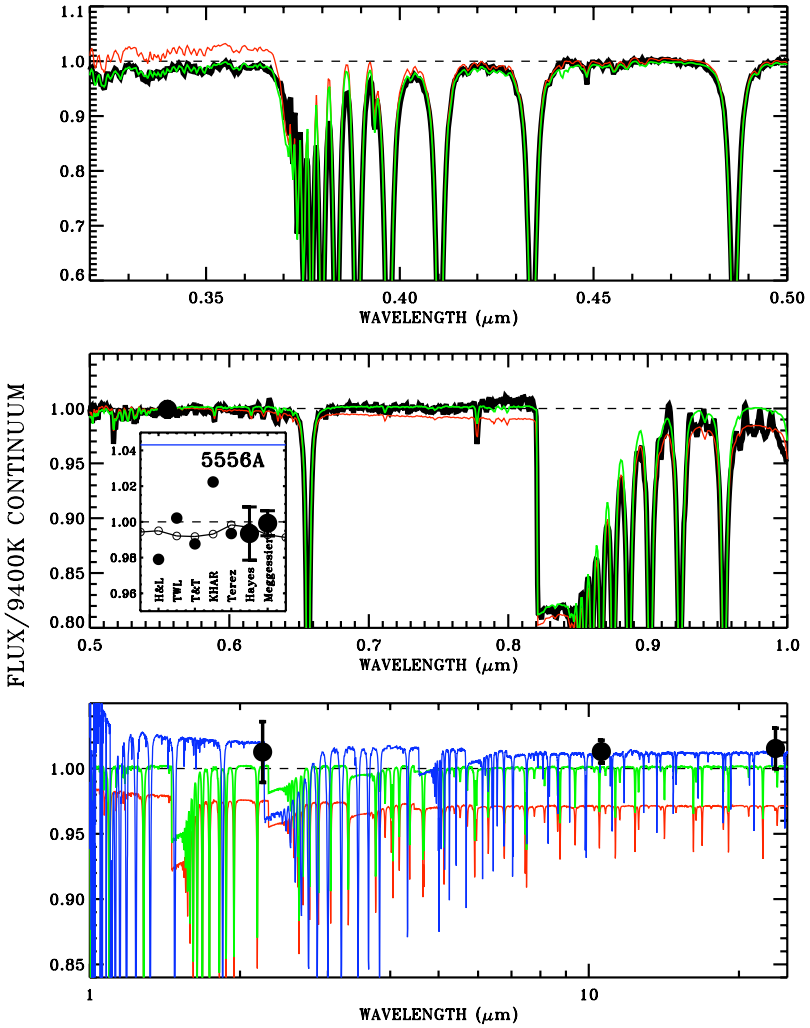

Figure 6. Ratio to the 9400 K model *continuum* level from 0.32 to 30 μm in three segments for the STIS measured fluxes (black), for the Kurucz 9400 K model (green), for the Kurucz 9550 K model (red), and for the Ri08 fluxes (blue). In the top two panels, the STIS spectrophotometry (black line) is mostly hidden under the green line. The three monochromatic, absolute IR-flux values of Ri08 appear on the blue Ri08 SED as filled circles with error bars. The Megessier flux value at 0.5556 μm is at unity, where the 0.7% error bar is inside the circle. The inset is a blowup of this 5556 Å region that shows this point along with the robust set of direct absolute flux measurements summarized by Hayes (1985). The five smaller filled circles are labeled per the Hayes nomenclature and are averaged to get his estimate of 3.44×10^{-9} erg cm^{-2} s^{-1} Å^{-1} at 5556 Å with its 1.5% error bar that is labeled as Hayes in the inset graph. Also in the inset, the nine small open circles connected with a line are the Hayes spectrophotometry on 25 Å centers from 5450 to 5650 Å.

standards. The ACCESS rocket program will establish a few fundamental flux standards in the brightness range of Sirius to $V \sim 9.5$ (Kaiser et al. 2007). A ground-based program, NIST STARS, is supported by lidar to measure the real-time atmospheric extinction and uses NIST calibrated detectors to establish an all-sky set of SEDs for standard stars with precisions of 0.5% (Zimmer et al. 2010).

6. SUMMARY

The IRAC calibration constants have been determined from a somewhat different set of flux standards than utilized by Re05. Small differences between our results and those of Re05 are explained by differences in the photometry extracted from the IRAC images. Our results are in agreement with the independent IRAC4 8 μm calibration of Ri08.

The robustness of our results suffer from a deficiency of G star flux standards, from the exclusion of two out of the four Re05 primary flux standards, and from offsets of up to 4σ between our set of WD stars and the cooler standards. To alleviate these deficiencies, more comparison standards with *HST* fluxes and better model grids are required to model the measured fluxes and establish SEDs to the *JWST* limit of 30 μm.

Support for this work was provided by NASA through the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. Thanks to A. Gianninas for supplying the 60,000 K SED with CNO that appears in Figure 3.

APPENDIX

HST CALIBRATION METHODOLOGY

A.1. Equations

The *HST* method of flux calibration for filter photometry does not involve color corrections or nominal wavelengths and is always defined in terms of the photon weighted mean flux over the bandpass in wavelength units

$$\langle F_\lambda \rangle = \frac{\int F_\lambda \lambda R \, d\lambda}{\int \lambda R \, d\lambda} = P_\lambda N_e(\text{pred}) \qquad \text{(A1)}$$

or in frequency units

$$\langle F_\nu \rangle = \frac{\int F_\nu \nu^{-1} R\, d\nu}{\int \nu^{-1} R\, d\nu} = P_\nu N_e(\mathrm{pred}) \quad \text{(A2)}$$

(Koornneef et al. 1986, Ri08), because detectors are generally photon detection devices, rather than total energy sensing bolometers. Some authors (e.g., Bessel et al. 1998; Cohen et al. 2003) define our product λR as their response function of the system.

Source independent calibration constants P are defined by dividing the mean flux by the predicted electrons s^{-1}, N_e(pred), in an infinite-radius aperture. If the predicted N_e(pred) from Equation (2) is substituted in Equations (A1) and (A2),

$$P_\lambda = \frac{hc}{A \int \lambda R\, d\lambda} \quad \text{(A3)}$$

$$P_\nu = \frac{h}{A \int \nu^{-1} R\, d\nu}. \quad \text{(A4)}$$

Because the *HST* standard flux units are normally per unit wavelength, the constant P_λ appears in the headers of *HST* photometric images with the keyword name *photflam*. For some instruments, e.g., NICMOS, P_ν with the keyword name *photfnu* is also included in the headers. The Re05 calibration constant in frequency units as per Equations (1) and (3) is

$$C' = \frac{F_{\nu_o} K}{N_e(\mathrm{pred})} = \frac{F_{\nu_o}}{N_e(\mathrm{pred})} \frac{\int \left(F_\nu / F_{\nu_o}\right)(\nu/\nu_o)^{-1} R\, d\nu}{\int (\nu/\nu_o)^{-2} R\, d\nu}. \quad \text{(A5)}$$

After substituting the definition of ν_o from Equation (4) with $\nu_o = c/\lambda_o$ and simplifying, the result is $\langle F_\nu \rangle / N_e$(pred), i.e., the *HST* and the Re05 methodologies produce exactly the same calibration constants $C' \equiv P_\nu$.

The *HST* calibration constants are normally derived from the source-independent Equations (A3) and (A4) after any required adjustments are made to the R estimated from the product of laboratory component quantum efficiency (QE) measurements. These adjustments are derived by making the measured N_e(obs) in an infinite aperture match the predicted N_e(pred) calculated from Equation (2). In practice, a radius of something like the 5.″5 for ACS is defined as "infinite" (Sirianni et al. 2005; Bohlin 2007b) and the primary pure hydrogen WDs G191B2B, GD71, and GD153 are the preferred standards used for F in Equation (2). Conversely for flux calibrated images as is the case for the *Spitzer* data, the measured N_e(obs) can be calculated from the measured mean-photometric flux as $N_e(\mathrm{obs}) = \langle F_\nu \rangle / P_\nu$. This reconciliation of laboratory component throughputs versus the truth of standard stars is achieved by adjusting the normalization of the filter throughput or even by changing the QE as function of wavelength for the detector or filter when sufficient information exists (e.g., de Marchi et al. 2004; Bohlin 2007b). Thus, information about individual component throughputs, such as the telescope or detector QE, may be inferred when there are multiple filters sampling the same wavelength regions.

To complement the above estimates of mean flux for stars imaged in a particular filter, an associated wavelength is often useful. In addition to the nominal wavelength λ_o of Re05, other common definitions are the *mean* and *effective* wavelengths.

$$\lambda_{\mathrm{mean}} = \frac{\int \lambda R\, d\lambda}{\int R\, d\lambda} \quad \text{(A6)}$$

$$\lambda_{\mathrm{eff}} = \frac{\int F_\lambda \lambda^2 R\, d\lambda}{\int F_\lambda \lambda R\, d\lambda}. \quad \text{(A7)}$$

Perhaps, most useful is the source independent *pivot-wavelength* λ_p and associated *pivot-frequency* ν_p, where $\lambda_p \nu_p = c$ and $\langle F_\lambda \rangle \lambda_p = \langle F_\nu \rangle \nu_p$.

$$\lambda_p = \sqrt{\frac{c\langle F_\nu \rangle}{\langle F_\lambda \rangle}} = \sqrt{\frac{\int \lambda R\, d\lambda}{\int \lambda^{-1} R\, d\lambda}}. \quad \text{(A8)}$$

These various measures of the associated wavelength for a filter are in Koornneef et al. (1986) along with a definition of the rms width of a filter. Appendix E of Ri08 contains our definition for mean wavelength but has an alternative formula for a bandwidth. The Ri08 definition of nominal wavelength is photon weighted as in Re05 but is an integral over wavelength in contrast to the Re05 integral over frequency. Also discussed in Ri08 is the concept of isophotal wavelength.

Having calculated the source independent *pivot-wavelength* λ_p, Equation (A8) provides a convenient formula for calculating P_λ from $P_\nu = C'$ values.

$$P_\lambda = \frac{c P_\nu}{\lambda_p^2}. \quad \text{(A9)}$$

A.2. Results

One complication in the comparison of the *HST* and Re05 flux calibration methodologies is that the above discussion is for the flux calibration of a point source, while the Spitzer images are calibrated in terms of surface brightness I. If a calibration constants C_I is defined so that multiplication by the original data numbers (DN/s) produces I, then

$$P_\nu \equiv C' = \frac{\Omega_{\mathrm{pix}} C_I}{G}, \quad \text{(A10)}$$

where Ω_{pix} is the size of a pixel in steradians, G is the gain in electrons/DN, and the DN is the unit of the instrumental signal. Another small complication is that the published values C of Re05 for the surface brightness calibration are adjusted for an absolute calibration based on a ten pixel radius for stellar photometry and require the correction to the infinite-aperture $C_I = C * f_{10}$. Values for f_{10} are listed in Re05 as 0.944, 0.937, 0.772, and 0.737 for IRAC channels 1, 2, 3, and 4, respectively, and may have large uncertainties because of complications caused by the small format and the internal scattering in the Si:As array. These values appear in Table 5 along with a selection of other calibration parameters, including *photflam* and *photfnu*, the *HST* style calibration constants. The top two rows of Table 5 are the published calibration constants from the *Spitzer* instrumental references in our Introduction. The row labeled "R-Corr." in Table 5 is the correction factor to the published *Spitzer* throughput values R that brings the N_e(pred) values calculated from Equation (2) into agreement with the measured $N_e(\mathrm{obs}) = \langle F_\nu \rangle / P_\nu$ for the weighted average of the eight A and G stars used for IRAC. The P_ν values are from Re05 and applying the tabulated R-Corr to the published throughput R curves makes the IRAC calibration self-consistent. Any change to an adopted calibration constant requires a corresponding change to the absolute level of that filter's R curve to maintain internal consistency. The R-Corr factors have poor precision for IRSB with its linearity problem and for MIPS with only three stars. Because the effective wavelength depends on the

Table 5
HST Calibration Methodology for Absolute Flux on the Re05 Scale

Item	Units	IRAC1	IRAC2	IRAC3	IRAC4	IRSB	MIPS24
C (Re05 10px)	MJy sr^{-1}/(DN s^{-1})	0.1088	0.1388	0.5952	0.2021	...	...
C_I	MJy sr^{-1}/(DN s^{-1})	0.1027	0.1301	0.4595	0.1489	0.0117[a]	0.0454
f_{10}		0.944	0.937	0.772	0.737	...	...
Gain	electrons/DN	3.30	3.71	3.80	3.80	...	5
$\Omega_{\rm pix}$	sr	3.498×10^{-11}	3.498×10^{-11}	3.498×10^{-11}	3.498×10^{-11}	7.893×10^{-11}	1.457×10^{-10}
$P_\nu = C'$	MJy/(e s^{-1})	1.089e-12	1.226e-12	4.230e-12	1.371e-12	9.235e-13	1.323e-12
P_λ	erg s^{-1} cm^{-2} Å^{-1}/(e s^{-1})	2.589e-20	1.819e-20	3.870e-20	6.613e-21	1.093e-21	7.026e-22
R-Corr		1.400	1.071	1.050	0.985	0.540	0.485
λ_o	μm	3.544	4.487	5.710	7.841	15.793	23.675
λ_p	μm	3.551	4.496	5.724	7.884	15.916	23.759
$\lambda_{\rm mean}$	μm	3.557	4.505	5.739	7.927	16.040	23.843
$\lambda_{\rm eff}$ (G191B2B)	μm	3.519	4.453	5.656	7.675	15.417	23.374
$\lambda_{\rm eff}$ (P330E)	μm	3.521	4.452	5.659	7.674	15.403	23.361

Note. [a] Units are per electron rather than DN, so a gain factor is not needed.

stellar SED, Table 5 includes examples for our hottest star G191B2B and a cooler star P330E; however, the differences are insignificant, because the slopes of the two SEDs are nearly the same in the IR. The flux weighting makes the $\lambda_{\rm eff}$ values smaller than the other measures of wavelength.

The IRSB relative response[10], R, has been divided by 3.58 to make the IRSB system throughput comparable to the other five modes and bring the values into agreement with Figure 6.1 in the IRS Data Handbook, where the peak QE is shown at a reasonable peak value of 0.81. Without this adjustment of 3.58, the value of 0.54 for the IRSB throughput correction in Table 5 would have been even smaller.

REFERENCES

Aldering, G., et al. 2004, arXiv:astro-ph/0405232
Asplund, M., Grevesse, N., & Sauval, A. J. 2005, in ASP Conf. Ser. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, ed. T. G. Barnes, III & F. N. Bash (San Francisco, CA: ASP), 25
Aufdenberg, J. P., et al. 2006, ApJ, 645, 664
Barstow, M. A., Good, S. A., Burleigh, M. R., Hubeny, I., Holberg, J. B., & Levan, A. J. 2003, MNRAS, 344, 562
Barstow, M. A., Hubeny, I., & Holberg, J. B. 1999, MNRAS, 307, 884B
Bessel, M. S., Castelli, F., & Plez, B. 1998, A&A, 333, 231
Bohlin, R. C. 2000, AJ, 120, 437
Bohlin, R. C. 2002, in 2002 HST Calibration Workshop, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore, MD: STScI), 115
Bohlin, R. C. 2007a, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor, MI: Sheridan Books), 315
Bohlin, R. C. 2007b, Instrument Science Report, ACS 2007-06 (Baltimore, MD: STScI)
Bohlin, R. C. 2010, AJ, 139, 1515 (B10)
Bohlin, R. C., & Cohen, M. 2008, AJ, 136, 1171
Bohlin, R. C., Dickinson, M. E., & Calzetti, D. 2001, AJ, 122, 2118
Bohlin, R. C., & Gilliland, R. L. 2004, AJ, 127, 3508
Bohlin, R. C., & Koester, D. 2008, AJ, 135, 1092
Castelli, F., & Kurucz, R. 2004, in IAU Symp. 210, Modeling of Stellar Atmospheres, ed. N. Piskunov, W. Weiss, & D. Gray (Cambridge: Cambridge Univ. Press), Poster A20 (CK04)
Cohen, M. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor, MI: Sheridan Books), 333
Cohen, M., Megeath, S. T., Hammersley, P. L., Martin-Luis, F., & Stauffer, J. 2003, AJ, 125, 2645
Cohen, M., Walker, R., Barlow, M., & Deacon, J. 1992a, AJ, 104, 1650
Cohen, M., Walker, R., & Witteborn, F. 1992b, AJ, 104, 2030 (CWW)
Cohen, M., Walker, R. G., Carter, B., Hammersley, P., Kidger, M., & Noguchi, K. 1999, AJ, 117, 1864
De Jong, R. 2006, Instrument Science Report, NICMOS 2006-03 (Baltimore, MD: STScI)
De Marchi, G., et al. 2004, Instrument Science Report, ACS 2004-08 (Baltimore, MD: STScI)
Diolaiti, E., Bendinelli, O., Bonaccini, D., Close, L., Currie, D., & Parmeggiani, G. 2000, A&AS, 147, 335
Engelbracht, C. W., et al. 2007, PASP, 119, 994
Engelke, C. W., Price, S. D., & Kraemer, K. E. 2010, AJ, 1919
Fazio, G. G., et al. 2004, ApJS, 154, 10
Finley, D. S., Koester, D., & Basri, G. 1997, ApJ, 488, 375
Gianninas, A., Bergeron, P., Dupuis, J., & Ruiz, M. T. 2010, ApJ, 720, 581
Gilliland, R. L., Goudfrooij, P., & Kimble, R. A. 1999, PASP, 111, 1009
Gilliland, R. L., & Rajan, A. 2011, Instrument Science Report, WFC3 2011-03 (Baltimore, MD: STScI)
Gordon, K. D., et al. 2005, PASP, 117, 503
Gustafsson, B., Edvardsson, B., Eriksson, K., Jorgensen, U. G., Nordlund, A., & Plez, B. 2008, A&A, 486, 951
Hayes, D. S. 1985, in IAU Symp. 111, Calibration of Fundamental Stellar Quantities, ed. D. S. Hayes, L. E. Pasinetti, & A. G. Davis Philip (Dordrecht: Reidel), 225
Hora, J. L., et al. 2008, PASP, 120, 1233
Houck, J., et al. 2004, ApJS, 154, 18
Hubeny, I., & Lanz, T. 1995, ApJ, 439, 875
Kaiser, M. E., Kruk, J., McCandliss Sahnow, D., Dixon, W., Bohlin, R., & Deustua, S. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor, MI: Sheridan Books), 361
Kent, S., et al. 2009, Astro2010: The Astronomy and Astrophysics Decadel Survey, Science White Papers, No. 155
Koornneef, J., Bohlin, R., Buser, R., Horne, K., & Turnshek, D. 1986, in Proc. IAU XIX General Assembly, Highlights of Astronomy Vol. 7, ed. J.-P. Swings (Washington, DC: AURA), 833
Maíz Apellániz, J. 2007, in ASP Conf. Ser. 364, The Future of Photometric, Spectrophotometric, and Polarimetric Standardization, ed. C. Sterken (Ann Arbor, MI: Sheridan Books), 227
Megessier, C. 1995, A&A, 296, 771
Peterson, D. M., et al. 2006, Nature, 440, 896
Price, S. D., Paxson, C., Engelke, C., & Murdock, T. L. 2004, AJ, 128, 889
Reach, W. T., et al. 2005, PASP, 117, 978 (Re05)
Rieke, G. H., et al. 2004, ApJS, 154, 25
Rieke, G. H., et al. 2005, ApJ, 620, 1010
Rieke, G. H., et al. 2008, AJ, 135, 2245 (Ri08)
Sirianni, M., et al. 2005, PASP, 117, 1049S
Smith, J. A., Tucker, D. L., Allam, S. S., & Rodgers, C. T. 2003, AJ, 126, 2037
Su, K. Y. L., et al. 2006, ApJ, 653, 675
Thuillier, G., Hersé, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., Simon, P. C., & Mandel, H. 2003, Sol. Phys., 214, 1
Tremblay, P. E., & Bergeron, P. 2009, ApJ, 696, 1755
Werner, M. W., et al. 2004, ApJS, 154, 1
Zimmer, P. C., et al. 2010, SPIE, 7735, 77358

[10] http://ssc.spitzer.caltech.edu/files/spitzer/bluePUtrans.txt